1 A POINT-OF-SALE COMMERCIAL TRANSACTION PROCESSING SYSTEM USING

2 ARTIFICIAL INTELLIGENCE ASSISTED BY HUMAN INTERVENTION

3

4 BACKGROUND OF THE INVENTION

5

6 1. Field of the Invention

7 This invention relates broadly to data processing systems for

8 commercial transactions. More particularly, this invention

9 relates to point-of-sale (POS) registers and systems for

10 communicating therewith to facilitate and expedite the ordering

11 and purchase of food items from a restaurant. Notably, the

12 invention utilizes artificial intelligence to at least partially

13 process transactions, and relies on human intervention where the

14 artificial intelligence is unable to complete the transactions.

15

16 2. State of the Art

17 The concept behind a fast food restaurant is the ability to

18 rapidly fulfill food orders placed by a customer at the

19 restaurant's order placement counter. In the current fast food

20 operation model, a plurality of point-of-sale (POS) registers are

21 located on the counter, and the registers are each operated by a

22 cashier behind the counter to enter a customer's order into the

23 register, for example via a keypad. The order is then

24 communicated, for example orally by the cashier, by printed

25 instructions, or by video display, to employees who prepare and

1   assemble the customer's order. In addition, the purchase price of

2   the order is totalled and the customer provides payment to the

3   cashier. Finally, the order is delivered to the customer, either

4   at the register or in an order pick-up queue.

5

6   The primary bottleneck to serving greater numbers of

7   customers is in the processing of the orders (i.e., order taken,

8   payment transaction, and order delivery). Research indicates that

9   profits for a fast food restaurant can be increased by decreasing

10   the transaction time for the orders, such that more orders can be

11   entered in a given time frame. However, using the order system

12   presently in place, order process time has been substantially

13   optimized. Training techniques for cashiers have been refined

14   over the years to arrive at the current techniques. While

15   one manner of increasing the ability to process orders would be to

16   provide additional point-of-sale registers on the order counter

17   and cashiers to operate the registers, counter space is limited.

18   Indeed, fast food restaurants are designed to provide a market

19   researched optimum split between order processing space (customer

20   waiting area and order counter, order fulfillment space (kitchen

21   and order preparation), and dining space. It would not be

22   desirable to disrupt the allocation of space within a fast food

23   restaurant.

24

1   A number of systems have been proposed and even attempted in

2   trials which are purported to increase order processing. For

3   example, U.S. Patent No. 5,235,509 to Mueller et al. discloses a

4   customer self-order system which displays menu items on a touch

5   screen and steps the customer through ordering from various food

6   categories: burgers, fries, salads, drinks, desserts, etc. U.S.

7   Patent No. 5,845,263 to Camaisa et al. discloses an interactive

8   visual order system which provides information in addition to menu

9   items and price to the customer. For example, the customer can

10   obtain information relating to method of preparation and

11   nutritional content, thereby allowing the customer to make a more

12   informed decision. None of these systems or other alternatives

13   has gained acceptance. It is believed that the failure of the

14   proposed systems all have a common drawback. In a fast food

15   environment, where lines of customers are frequently encountered,

16   some customers may be intimidated or confused by the unfamiliar

17   systems and require employee assistance, which slows down the

18   entire system. An additional drawback to the proposed systems is

19   their inability to effectively promote sales with the degree of

20   success provided by a human cashier. Customers all know the

21   ubiquitous phrase "do you want fries with your order". The phrase

22   is used so commonly because it effectively increases sales. In

23   addition, the current trend to move customers to an 'upsized'

24   order of french fries or soft drink also substantially increases

25   the sales at a restaurant, and any new system must be able to

1 provide such promotional features as effectively as a human

2 cashier. Otherwise, the systems will not gain favor by the

3 restaurant operators and will not be utilized.

4

5   SUMMARY OF THE INVENTION

6

7   It is therefore an object of the invention to provide a

8 system which can process a great number of fast food orders.

9

10   It is another object of the invention to provide an order

11 system which provides a familiar order experience to the customer.

12

13   It is a further object of the invention to provide a fast

14 food ordering system which optimizes the use of order processing

15 space in a fast food restaurant.

16

17   It is an additional object of the invention to provide a fast

18 food ordering system in which the numbers of behind-the-counter

19 cashiers may be reduced or even eliminated, thereby providing

20 additional room for order fulfillment such that additional orders

21 may be processed.

22

23   Another object of the invention is to provide a fast food

24 ordering system which does not require customer 'training' and

1 provides a relatively seamless experience for the customer

2 relative to conventional fast food ordering.

3

4   A further object of the invention is to provide a fast food

5 ordering system which effectively promotes products in a manner to

6 increase sales for the restaurant.

7

8   It is yet another object of the invention to provide a point-

9 of-sale commercial transaction processing system which is

10 adaptable for use in a variety of commercial industries and

11 establishments.

12

13   In accord with these objects, which will be discussed in

14 detail below, a point-of-sale commercial transaction processing

15 system, particularly suitable for the fast food industry, is

16 provided. The transaction processing system utilizes (1) a

17 customer interaction terminal (CIT) having a video display, an

18 audio speaker, a microphone, and preferably a printer, (2) a

19 computer system coupled or integral with the customer interaction

20 terminal (CIT) and running artificial intelligence routines to

21 process or pre-process verbal requests provided into the

22 microphone of the customer interaction terminal, and (3) a human-

23 controlled response system which completes, corrects or verifies

24 requests that cannot be satisfactorily completed by the artificial

25 intelligence routines alone. The human-controlled response system

1   is preferably in communication with the customer interaction

2   terminal (CIT) (and the customer) via a high speed voice over

3   internet protocol (VoIP) or data connection.

4

5   According to a preferred embodiment of the invention, the

6   computer system presents on the customer interaction terminal

7   (CIT) a graphic image of a virtual cashier which is programmed to

8   interact graphically and through audio with the customer in a

9   manner to which the customer is accustomed from prior experience

10 with human cashiers in conventional fast food restaurants. That

11 is, the virtual cashier preferably includes an image of a face of

12 cashier which auditorily greets, engages, and prompts the customer

13 to verbally provide the fast food order to the virtual cashier

14 (Hello. Please tell me your order.). The customers verbal orders

15 are received by the microphone of the CIT and transmitted to the

16 computer system where they are processed. As the virtual cashier

17 image is computer generated, the face and other features of the

18 cashier may be human-like or whimsical, and may even be

19 representative of a mascot of the restaurant.

20

21   The artificial intelligence routines of the computer system

22 are preferably adapted to process the verbal orders such that a

23 complete fast food order (menu item selection, special preparation

24 requests, eat in or take out, etc.) can be processed. The

25 complete order may require multiple interactions between the

- 6 -

7

1   customer and virtual cashier; i.e., after the customer orders a

2   sandwich, the virtual cashier can engage the customer and ask

3   whether the customer would like a soft drink and, if so, which

4   size. Furthermore, according to a preferred aspect of the

5   invention, the routines in the computer system which operate the

6   virtual cashier are adapted to follow techniques which are shown

7   to increase restaurant sales. For example, the virtual cashier

8   can ask whether the customer would like french fries with an

9   order, or whether for a nominal additional sum the customer would

10 prefer to upsize the french fries and drink order. In addition,

11 the virtual cashier can promote special offers and provide

12 advertisements.

13

14 It is recognized that current state of the art artificial

15 intelligence alone may not be sufficient to satisfactorily

16 complete all fast food orders. As such, according to a preferred

17 aspect of the invention, when the computer system is unable to

18 satisfactorily complete a fast food order via the human customer-

19 artificial intelligence virtual cashier interaction, or at any

20 time upon customer request, a human-controlled response system,

21 preferably located off-site of the CIT, is employed to complete,

22 correct or verify the order by interaction with the customer via

23 the CIT. The interaction of the human-controlled response system

24 is through the graphics and audio of the CIT and is preferably

25 indistinguishable to the customer relative to interaction with the

1   artificial intelligence routines. That is, the customer is

2   preferably unaware of any shortcoming in the AI processing and

3   perceives the order interaction as one continuous interaction even

4   if the response system is utilized within an order. The

5   availability of human intervention permits the use of artificial

6   intelligence even as the state of the art artificial intelligence

7   may not yet be ripe for use in all fast food order transactions.

8

9   Once an order has been completely processed, payment may be

10   made at the CIT, using a debit or credit card or cash, and the

11   order is sent to order fulfillment employees who prepare the

12   order. The customer is also directed to a pick-up location and

13   may be given an order number corresponding to his or her order.

14

15   It will be recognized that the above described system

16   eliminates the need and space required for traditional human

17   cashiers and, therefore, a greater amount of the order processing

18   space may be devoted to CITs. In addition, CITs may be placed on

19   tables, on walls, at kiosks, at drive-through locations, in

20   portable devices, and at other locations. Furthermore, as the

21   CITs preferably display a face and provide a spoken dialogue with

22   the customer, the customer does not require any particular

23   training to use the system; i.e., use of the system of the

24   invention provides a substantially seamless experience, in terms

25   of ordering, from conventional fast food ordering experiences.

- 8 -

9

1 The customer interacts in the same manner as he or she has

2 previously with human cashiers. Moreover, the system, while easy

3 to use for the customer, provides substantial novelty which

4 attracts and retains customers.

5

6   Additional objects and advantages of the invention will

7 become apparent to those skilled in the art upon reference to the

8 detailed description taken in conjunction with the provided

9 figures.

10

11   BRIEF DESCRIPTION OF THE DRAWINGS

12

13   Fig. 1 is a schematic diagram of a point-of-sale commercial

14 transaction processing system according to the invention;

15

16   Fig. 2 is a flow chart of a first embodiment of implementing

17 the point-of-sale commercial transaction processing system of the

18 invention; and

19

20   Fig. 3 is a flow chart of a second embodiment of implementing

21 the point-of-sale commercial transaction processing system of the

22 invention.

23

10

1   DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

2

3   Turning now to Fig. 1, a point-of-sale commercial transaction

4   processing system 10 particularly suited for a fast food

5   establishment is provided. The transaction processing system

6   includes a customer interaction terminal (CIT) 12, a computer

7   system 14 coupled to the CIT, and a human-controlled response

8   system 16 in communication with the computer system 14.

9

10   The CIT 12 includes a video display 20, an audio speaker 22,

11   a microphone 24, and optionally a video camera 26. In addition,

12   the CIT 12 also preferably includes a printer 28, a debit/credit

13   card reader 30, and a bill and/or coin currency reader 32 as well

14   as a change dispenser 34. The CIT may also include an activation

15   button 36, such as a 'push-to-talk' button, and may further

16   include a sensor 37, e.g., an infrared or sonic sensor, which

17   senses when a customer is located in an "ordering" position

18   relative to the CIT. As an optional alternative, the video camera

19   26 may function as the sensor. The CIT 12 is located in a fast

20   food restaurant. The CIT 12 may be placed on a counter, in a

21   kiosk, on a wall, at a dining table, along a take-out drive-

22   through route, in a portable device which may be transported along

23   a drive-through route, or in any other suitable location within or

24   relative to a fast food restaurant enabling customer interaction

25   with the CIT.

11

1   The computer system 14 is coupled to or integral with one or

2   more CITs and is adapted to receive input from the CITs 12 (via

3   the microphone 24 and optional video camera 26) and provide output

4   to the display 20, audio speaker 22, and printer 28 of the CIT.

5   That is, the CIT 12 is under the control of the computer system

6   14.   In addition, the computer system 14 preferably includes a

7   memory adapted to record the audio and optionally the video

8   portion of a current interaction between a customer and a CIT.

9   While multiple CITs 12 may be coupled to a single computer system

10   14 (two CITs 12 being shown in solid lines in Fig. 1), for

11   clarity, the invention will be described with respect to a single

12   CIT 12 being coupled to the computer system 14.   The computer

13   system 14 has software adapted to permit each CIT 12 to 'interact'

14   with a customer and process (via artificial intelligence routines)

15   customer orders spoken into the microphone 22, as described in

16   more detail below, and a microprocessor adapted to run the

17   software.

18

19   The human-controlled response system, e.g., a call center 16,

20   is connected to the computer system 14 (or multiple computer

21   systems 14, each, in turn, coupled to one or more CITs 12).   The

22   call center 16 is preferably located on different premises than

23   the CIT 12 and computer system 14, and more preferably located in

24   a country or region having a relatively lower labor cost than the

25   country or region in which the CIT is located.   A number of human

- 11 -

12

1   operators 40 work at the call center, and each operator is

2   provided with an audio speaker 42, a microphone 44, and a display

3   46. The audio speaker 42 is adapted to reproduce for the operator

4   40 sounds (words) spoken into the microphone 24 of the CIT and/or

5   recorded by the computer system 14, the microphone 44 is adapted

6   to permit the operator to provide spoken messages to the customer

7   via the speaker 22 of the CIT, and the display 46 permits the

8   operator to see the customer's order, and preferably displays the

9   same images shown on the display 20 of the CIT.

10

11   Referring to Figs. 1 and 2, the software permitting the CIT

12   12 to 'interact' with customers includes a graphic image of a

13   virtual cashier 38 programmed to interact graphically and through

14   audio via the microphone 22 and speakers 24 with the customer

15   interfacing with the CIT 12. The images of the virtual cashier

16   are preferably computer generated and, as such, the face and other

17   features of the cashier may be human-like, animal-like, or

18   whimsical in nature, and may even be representative of a mascot of

19   the restaurant (e.g., Ronald McDonald) or characters in a movie or

20   television show. Human-like features may be representative of

21   celebrities. The interaction is preferably performed in a manner

22   similar to that which the customer is accustomed from prior

23   experience with human cashiers in conventional fast food

24   restaurants. That is, the virtual cashier preferably displays

25   images of a face of cashier and auditorily greets, engages, and

- 12 -

13

1 prompts at 100 the customer to verbally provide a fast food order

2 to the virtual cashier (e.g., "Hello. Please place your order

3 with me."). The customer's verbal orders are received by the

4 microphone 22 of the CIT and transmitted to the computer system 14

5 where they are processed, as described below.

6

7 According to a preferred embodiment of the order processing

8 software, when a customer order is verbally provided into the

9 microphone 22 of the CIT at 102, the order is provided to the

10 computer system 14 and the artificial intelligence routines are

11 adapted to process at 104 the customer's order in real time. That

12 is, the artificial intelligence (AI) routines are adapted to parse

13 from the orders the necessary content to determine what the

14 customer wants to order.

15

16 The ability of the AI routines to satisfactorily process

17 customer orders at 104 depends on the amount of variability

18 present in the process; i.e., the extent to which the vocabulary

19 and the grammar used in the interaction varies from one customer

20 to another. The AI routines are preferably optimized based on

21 conditioning data collected from conversation over a period of

22 time (e.g., a few days) at a conventional cashier point of sale

23 terminal and examined for recurring patterns, and then used to

24 train the AI routines. AI routine training and optimization is

25 preferably performed on a continual basis, with reports of

14

1 misunderstood communications regularly analyzed and used to

2 improve the performance of the system.

3

4   An important issue for automating customer interaction with

5 the CIT is to distinguish between customer-CIT interaction speech

6 and other speech, e.g., utterances by the customer to other people

7 in the vicinity, or even "talking to oneself". One simple

8 approach to overcome this difficulty is to use a push-to-talk

9 button 36, as described in J. Gustafson, N. Lindberg, and M.

10 Lundeberg, "The August Spoken Dialogue System", *Proceedings of*

11 *Eurospeech '99* (1999). Another preferred approach is to use the

12 optional video camera 26 to track the customer's head orientation

13 and gaze (head tracking), and only respond to utterances made when

14 the user is looking directly at the CIT. The problem of pose

15 recognition (i.e., recognizing, from a camera image, whether a

16 person's face is oriented towards the camera) is not very

17 difficult. For example, standard machine learning techniques can

18 be used. First, a training corpus of faces is constructed, with

19 examples of faces looking at the camera and faces looking

20 elsewhere. Second, the system is trained to learn an algorithm

21 which distinguishes between the two. Finally, the resulting

22 classifier is applied to new faces. Reasonable success has been

23 achieved on the pose recognition task using neural networks. T.

24 Mitchell, *Machine Learning*, McGraw Hill (1997). More modern

25 classifiers, such as support vector machines can be used to

1 achieve even higher accuracy. C. Burges, "A Tutorial on Support

2 Vector Machines for Pattern Recognition", *Data Mining and*

3 *Knowledge Discovery*, Vol. 2, Number 2, p. 121-167 (1998). There

4 are also approaches based on template matching which may be used.

5 B. Scassellati, "Finding Eyes and Faces with a Foveated Vision

6 System", *Proc. 15th National Conference on Artificial Intelligence*

7 *(AAAI-98)*, AAAI Press (1998).

8

9   Tracking the customer's head orientation is user-friendly,

10 especially in drive-through settings, where the customer would

11 otherwise have to extend his or her hand from the car to push the

12 button 36. However, it is also more prone to error, particularly

13 for nonstandard facial configurations (e.g., men with a heavy

14 beard or people wearing a baseball cap). When the head tracking

15 approach is used, it is preferable that the CIT 12 indicate to the

16 customer when the system interprets that the customer's spoken

17 words are aimed at the system. Where the interaction is based on

18 animated characters, the indicator can be entertaining. For

19 example, if the system is not listening to the spoken words of the

20 customer, the character can pretend to be sleeping. Then, when

21 the customer is communicating with the system, but the system

22 fails to recognize the communication, the manual push-to-talk

23 button can be used as a fall-back.

24

76

1   Current speech recognition systems are effective in one of

2   two modes: either there is a single speaker, for which the system

3   can be individually trained to understand a relatively large

4   vocabulary; or there can be multiple speakers for which current

5   systems can recognize a limited vocabulary. For the system of the

6   invention, the vocabulary required is likely to be quite limited,

7   thereby making high-accuracy speech recognition feasible for use

8   by multiple users. For example, in another application of a

9   restricted-domain automated dialogue system, a vocabulary of 500

10   words was sufficient. See Gustafson et al. (1999).

11

12   Currently existing on the market are several high-quality

13   commercial speech recognition systems: Dragon's

14   NaturallySpeaking™, Kurzweil Applied Intelligence's L&H Voice

15   Xpress™, and IBM's ViaVoice™. These speech recognition systems

16   are primarily designed for single speaker, large vocabulary

17   settings. However, the systems may be modified for use in a

18   multiple speaker, limited vocabulary setting. Another option is

19   to utilize a speech recognition tool kit with a developer's

20   application programming interface (API) that allows the vocabulary

21   and other aspects to be tailored to the fast food ordering process

22   (See http://www.speech.cs.cmu.edu/comp.speech/FAQ.Packages.html

23   for links to available packages.) In addition, an existing public

24   domain speech recognition system may be adapted, or a system may

17

1 be implemented from scratch. *See* Gustafson et al. (1999) for

2 details on how this can be done.

3

4   An important issue is the ability of the AI routines of the

5 computer system 14 to recognize when the result of the speech

6 recognition is correct and, when it is incorrect, to cause the

7 system to ask the customer for a clarification or cause

8 intervention by a human operator, as described further below.

9 Several approaches can be used for estimating this confidence.

10 All current speech recognition systems use an underlying

11 probabilistic model, so they can be adapted to output the

12 probability of the acoustic signal given the recognized words. In

13 other words, this number estimates how likely this particular word

14 sequence is to have generated the acoustic signal heard. If this

15 number is low, this is an indicator of possibly faulty

16 recognition. A possible improvement to this approach is to also

17 generate a second probability of the acoustic signal given a

18 syllable-based recognition system that does not try to match

19 words. *See* Gustafson et al. (1999). If the second probability is

20 substantially higher than the first, then the utterance contained

21 words outside the vocabulary of the speech recognition system's

22 lexicon that the system "forced" into words in the lexicon.

23

24   As the AI routines processes the words of the customer's

25 order, the task is to recognize the customer's request at a level

18

1   sufficient to process his order. There are several approaches to

2   this task, of increasing complexity on the one hand, but probably

3   higher accuracy rates on the other. The simplest approach is to

4   use no semantic processing, just recognition of basic menu items.

5   In this case, the analysis is a direct result of the speech

6   recognition. If one recognizes the words "three" and

7   "SuperBurgers" in an utterance, this is interpreted as an order of

8   three SuperBurgers. This approach will work for simple menu

9   orders, but may be too limited to many cases, as it may not be

10   able to deal with any extensions such as "without pickles", "extra

11   mustard", etc.

12

13   A second level is to generate a corpus of utterances that are

14   likely to occur in a customer-CIT interaction. One can then

15   compare the customer utterance to others in the database, and find

16   the closest match. This is the approach discussed in Gustafson et

17   al. (1999). A semantic interpretation, i.e., a mapping between

18   the sentence structure and an order form, can then be manually

19   constructed for each template sentence. The extent to which this

20   approach can be successful depends, as discussed above, on the

21   variability of utterances that occur.

22

23   The next level is to actually parse the sentences using some

24   type of grammar. There has been substantial improvement in the

25   last five years in parsing (See C. Manning and H. Schütze,

1 *Foundations of Statistical Natural Language Processing*, MIT Press

2 (1999)), including a recent account of parsing a large corpus of

3 speech utterances. See E. Charniak, "The Statistical Natural

4 Language Processing Revolution", colloquium talk given at Stanford

5 University, April 26, 2000 (see http://robotics.stanford.edu/ba-

6 colloquium/previous/spring00/abst-charniak.html. The parsing uses

7 a grammar, preferably learned automatically from a corpus of

8 utterances parsed manually. As discussed in L. Bell and J.

9 Gustafson, "Interaction with an Animated Agent in a Spoken

10 Dialogue System", *Proceedings of Eurospeech '99* (1999), the number

11 of grammatical variations in automated dialogue systems is usually

12 quite small, and the grammar is quite simple. Again, it is

13 easiest to manually provide a semantic interpretation for each

14 grammatical structure, as above.

15

16 A difficult problem is the treatment of anaphoric references,

17 of the form: "actually, cancel that and give me two orders

18 instead". It is difficult to relate the words "that" and

19 "instead" to particular items used in previous utterances. There

20 has been some success in automatically clarifying anaphoric

21 references (*See* E. Charniak, N. Ge, and J. Hale, "A Statistical

22 Approach to Anaphora Resolution", *Proceedings of the Sixth*

23 *Workshop on Very Large Corpora* (1998)), but this is still a

24 difficult problem. As such, the AI routines are preferably

25 adapted to clarify the references by having the CIT ask the

20

1 customer "what do you mean?", or by causing intervention by a

2 human operator 40, discussed below.

3

4   If the AI routines are able to recognize and parse the speech

5 of the customer (as preferably determined via a probabilistic

6 calculation) such that individual menu items ordered are properly

7 added to the customer's order list (the transaction) after each

8 menu item is ordered, the CIT is updated at 108 (and 42 in Fig. 1)

9 to affirm order recognition. The update preferably includes one

10 or more of three subtasks: text generation, voice generation, and

11 animation generation. The text is preferably displayed in a

12 predetermined set of grammatical forms, filled in with the details

13 of the customer's transaction. Voice generation to interact with

14 the customer may be based on automated speech synthesis. However,

15 given the limited number of words that the CIT would need to

16 reproduce, it is preferable that CIT speech to a customer be

17 provided using pre-recorded words. Known smoothing techniques are

18 preferably used to provide a natural sounding transition between

19 the reproduced pre-recorded words. The virtual cashier's face is

20 animated to correspond to the words being 'spoken' by the virtual

21 cashier. This can be done by one of two approaches. The simpler

22 approach is a 'manual' approach in which, for human characters, a

23 human actor prerecords all the words which may need to be spoken

24 and, using standard morphing techniques, the transitions between

25 words are smoothed. For cartoon characters, each word can be

21

1 animated and the same morphing technique can be used for the

2 transition.

3

4   A more complex approach permits more sophisticated

5 interactions. In this approach, a computer-generated character is

6 animated based on an actor's rendition of the same word. The

7 actor says a word with certain markers on his face, capturing the

8 main articulation points. The articulation points are then mapped

9 onto corresponding points for the cartoon character, allowing the

10 character's animation to mimic the actor's expression. Again,

11 simple morphing can be used to deal with the inter-word

12 transitions. See, for example, F. Pighin, J. Hecker, D.

13 Lischinksi, R. Szeliski, and D. Salesin, "Synthesizing Realistic

14 Facial Expressions from Photographs", *Proceedings of Siggraph*

15 (1998), and M. Brand, "Voice Puppetry", *Proceedings of Siggraph*

16 *'99* (1999). In addition, basic emotional affect and other

17 interactive changes to the facial animation can be incorporated.

18 For example, the eyes of an animated character can follow the

19 customer using feedback from the video camera 26. See Gustafson

20 et al. (1999) and Pighin et al. (1998).

21

22   In addition, the CIT 12 may prompt the customer via computer-

23 generated voice or displayed text to add other items to the menu

24 list, and a complete order may require multiple interactions

25 between the customer and virtual cashier 38; i.e., after the

- 21 -

1 customer orders a sandwich, if the customer does not on his or her

2 own add additional menu items within a predetermined time period,

3 e.g. two second, the virtual cashier engages the customer and asks

4 whether the customer would like a soft drink and, if so, which

5 size. Furthermore, according to a preferred aspect of the

6 invention, the routines in the computer system which operate the

7 virtual cashier are adapted to follow additional techniques which

8 are shown to increase restaurant sales. For example, even after

9 the customer indicates that his or her order is complete, the

10 virtual cashier preferably asks whether the customer would like

11 french fries with an order which does not already include french

12 fries, or whether for a nominal additional sum the customer would

13 prefer to upsize the french fries and drink order, e.g., from

14 medium to large. Furthermore, at any time during the order

15 process, the virtual cashier can promote special offers and

16 provide advertisements for products in the restaurant

17 establishment or for products from outside establishments.

18 Additional menu items orders are processed at 104 and the CIT is

19 updated at 108 until the customer indicates at 110 that the order

20 is complete.

21

22   If at any time during the customer's order placement there is

23 a problem with the order processing at 106, as preferably

24 determined via a probabilistic calculation, (and optionally at any

25 time upon request by the customer, e.g., by pressing a button or

- 22 -

1 by verbal request), a network connection is created at 112 between

2 the CIT 12 (and computer system 14) and the call center 16. The

3 network connection may be a high speed voice over internet

4 protocol (VoIP) connection permitting the transmission of the

5 customer's voice order quickly and inexpensively to an operator 40

6 at the off-site call center 16. Additionally or alternatively,

7 the connection may be a high speed data connection, and the

8 customer's recorded verbal order is sent from the memory of the

9 computer system 14 to the audio speaker 42 directed at the

10 operator 40. The operator 40 is able to correct, verify or

11 complete the customer menu orders at 114. As the operator makes

12 the required changes or additions, the CIT is updated at 116 to

13 indicate the changes and additions and provide feedback to the

14 customer. Whether the AI routines or an operator is interacting

15 with the customer, according to the preferred embodiment of the

16 invention, it is desirable that the customer receive the same

17 manner of interaction so that the customer is unaware when an

18 operator 40 has intervened. As such, instructions by the operator

19 40 to the CIT 12, at 116, preferably result in the same type of

20 CIT updating (text, speech, and animation) as when the AI routines

21 alone interface with the customer and, until the order is complete

22 at 118, the customer continues his or her food order by speaking

23 to and otherwise interacting with the virtual cashier 38 on the

24 CIT at 120. The operator preferably interacts with the CIT and

25 the customer by inputting keyboard commands, mouse, or voice

- 23 -

24

1 commands which cause a preprogrammed automated update responses at

2 the CIT. If the operator needs to respond outside the capability

3 of the preprogrammed responses, the operator preferably speaks

4 into the microphone 44 and the speech is converted to text by

5 voice recognition. The recognized speech is filtered to remove

6 unwanted accents and words and to provide smoothing, and data

7 corresponding to the speech is sent to the computer and then

8 synthesized by the CIT or used to trigger recorded words in memory

9 of the CIT. According to the first embodiment of the invention,

10 once the connection is made with the call center at 112, the

11 operator 40 is utilized to complete the order with the customer

12 without reversion to the AI routines.

13

14 Once the order is complete at 118, the customer is prompted

15 at 122 for payment which is preferably made at the CIT. Payment

16 is made at 124 using a debit card or credit card in conjunction

17 with the card reader 30, or with cash in conjunction with the bill

18 reader 32 and change dispenser 34. After payment is made, the CIT

19 prints at 126 with the printer 28 a receipt for the customer

20 indicating the details of the order as well as an order number,

21 and the virtual cashier directs the customer to proceed to an

22 order pick-up area. In conjunction with order payment and receipt

23 printout, the order is sent at 128 to order fulfillment employees

24 (kitchen staff and order assembly personnel) who prepare the

25 order. The orders are packaged with the respective order number

1   and, once complete, the customer is provided at 130 with the

2   customer's corresponding order.

3

4   Turning now to Figs. 1 and 3, a flow chart for a second

5   preferred embodiment of the invention is shown. The second

6   embodiment is substantially similar to the first embodiment, with

7   the following differences. A CIT greeting is provided at 200

8   which, rather than prompts the customer to place an immediate

9   order (as in the first embodiment), requests whether the customer

10   would like to place an order, e.g., "Hello, would you like to

11   place an order." This request is intended to cause an initial

12   "Yes" response or other CIT-customer interaction from the

13   customer, at 202, prior to order placement and provide a short

14   delay prior to order entry which is sufficient for establishment

15   at 204 of a connection to the call center 16. Alternatively, the

16   connection may be made upon indication by a sensor 37 which senses

17   the presence of customer ready to place an order. As yet another

18   alternative, a constant connection may be maintained between the

19   CIT 12 and the call center 16 and the CIT greeting may be intended

20   to cause immediate order placement by the customer.

21

22   In either approach, the customer then interacts at 206 with

23   the CIT 12 in real-time, verbally ordering food. The AI routines

24   in the computer process the interaction at 208 to parse and

25   identify the elements the food order. Assuming there is no

26

1   problem with the processing at 210, after each menu item is

2   ordered, the CIT is updated at 212, and the AI routines continue

3   to process the order until the order is complete at 214. However,

4   if there is a problem at 210 during any of the AI processing, the

5   order is assigned to an operator 40 at the call center 16, and the

6   operator corrects the order at 216, and the CIT is then updated at

7   212. According to the second embodiment of the invention, if the

8   customer order is incomplete at 214, the AI routines are again

9   given responsibility at 208 for processing the interaction between

10   the customer and CIT at 206 and maintains control absent another

11   processing problem at 208. This is in contrast to the first

12   embodiment, where after the occurrence of a processing problem an

13   operator is given responsibility for not only correcting the

14   problem but completing the order.

15

16   Once the order is complete at 214, the steps of prompting the

17   customer for payment through providing the customer with the order

18   (that is, steps 222-230) are the same as the analogous steps in

19   the first embodiment (steps 122-130).

20

21   While the above described transaction processing system is

22   optimized for use within a fixed location, such as the order

23   processing space of a fast food restaurant, it will be appreciated

24   that the CIT may be optimized for drive-through use and adapted to

25   be handed to the customer or taken from a station at the beginning

- 26 -

1   of a drive-through route. In order to avoid driving accidents due

2   to diverted attention while using the portable CIT, the portable

3   CIT preferably includes an accelerometer, which allows the unit,

4   as well as an operator at the call center, to know whether the

5   customer's vehicle is in motion. The CIT is optionally programmed

6   to not interact with the customer while the vehicle is in motion.

7   For example, the portable CIT can repeat a message to ask the

8   customer to continue the ordering process once the vehicle is

9   stopped. The portable CIT is preferably formed to fit within a

10 standard cup holder found in most cars, and the top of the

11 portable CIT is preferably provided with a small display screen

12 which preferably alternately displays the virtual cashier and a

13 screen that lists the menu items being ordered. The portable CIT

14 preferably contains a debit/credit card reader to facilitate and

15 expedite payment. The portable CIT optionally includes a

16 compartment in which the customer can place paper and coin

17 currency. The portable CIT is returned to a restaurant employee

18 at the time of order pickup. If the customer pays with cash, the

19 employee will remove the cash from the compartment in the CIT and

20 give change to the customer. Finally the customer receives the

21 food ordered.

22

23   Another option is for the CIT to be an all audio-based

24 device, without a display component. The benefits of such a CIT

25 is that the hardware and software for the device are cheaper and

28

1 more reliable. One exemplar all audio-based CIT eliminates the

2 speaker and preferably includes written instructions directing the

3 customer to tune a car radio to a particular frequency. The CIT

4 then broadcasts the virtual cashier's voice into the car through

5 the car's radio and speaker system. This may be done by adapting

6 the CIT such that when it is placed near the car radio, it

7 automatically sends the audio signal over the car's radio system.

8 Transmission of audio signals through the radio in this manner is

9 known for common audio devices such as MP3 players. In addition,

10 rather than include a microphone, the words spoken by the customer

11 are received by means of a laser incident on the windshield or

12 driver side window of the car. The laser detects the vibration of

13 the glass caused by the customer's spoken words and then reconvert

14 this vibration signal back into an audio signal. This technology,

15 developed for espionage purposes, is now widely available. The

16 advantage of this approach is that it minimizes the need for extra

17 hardware to be produced and then put at risk by placing it into

18 the hands of the customer where it potentially may be damaged or

19 stolen.

20

21 If multiple CITs are distributed to drive-through customers,

22 it is preferable that each is linked to a central server in the

23 restaurant by wireless networking technology, e.g., such as the

24 Bluetooth™ standard. In addition, it is preferred that the

25 portable CITs be used in conjunction with a system which prevents

29

1 or inhibits accidental or purposeful removal of the CIT from the

2 restaurant property. As such, when the unit is removed from the

3 restaurant property, the unit is preferably adapted to make an

4 alarm sound and warn the customer to return the CIT unit. The

5 restaurant staff is likewise alerted and a digital or film

6 photograph is preferably taken of the car (including the license

7 plate) to aid in law enforcement action recovery. The portable

8 CIT preferably informs the customer that a picture has been taken

9 of their car and instructs the customer to return the unit to the

10 restaurant. The CIT may also send out a tracking signal, e.g.,

11 GPS coordinates, permitting the CIT to be located.

12

13   It will be recognized that the above described systems and

14 methods eliminate the need and space required for traditional

15 human cashiers and, therefore, provide a greater amount of the

16 order processing space for CITs. Furthermore, as the CITs

17 preferably display a face and provide a spoken dialogue with the

18 customer, the customer does not require any particular training to

19 use the system; i.e., use of the system of the invention provides

20 a substantially seamless experience, in terms of ordering, from

21 conventional fast food ordering experiences. The customer

22 interacts in the same manner as he or she has previously with

23 human cashiers. Moreover, the system, while easy to use for the

24 customer, provides substantial novelty which attracts and retains

25 customers.

1   There have been described and illustrated herein several

2   embodiments of a point-of-sale commercial transaction processing

3   system, and one particularly suited for use in a fast food

4   restaurant. While particular embodiments of the invention have

5   been described, it is not intended that the invention be limited

6   thereto, as it is intended that the invention be as broad in scope

7   as the art will allow and that the specification be read likewise.

8   Thus, while particular elements of the CIT have been disclosed, it

9   will be appreciated that other elements may be included or

10   removed, provided that the CIT is capable of permitting verbal

11   input from the customer which can then be at least partially

12   processed by AI routines in a computer. Furthermore, while in the

13   first embodiment, the operator once given control of a portion of

14   a customer order retains control of the order, it will be

15   appreciated that the system can be operated permit the AI routines

16   to regain control of an order. In addition, while particular

17   orders of the method of the invention have been shown and

18   described with respect to the flow charts, it will be appreciated

19   that another order may be used, and that the two flow charts are

20   exemplary. Also, while the transaction processing system has been

21   described with respect to the operations of a fast food

22   restaurant, it will be appreciated that the system may be used in

23   other industries which have conventionally used a point-of-sale

24   register. By way of example, and not by way of limitation, the

25   system is suitable for use in the rental car industry and the

31

1 purchase of movie/theater tickets. Furthermore, while the display

2 is shown with a virtual cashier and details of an order, it will

3 be appreciated that the display can display advertising (of the

4 establishment in which it is being used, or of another

5 establishment, and promotions of the establishment). Such

6 displays of advertising and promotions can occur during an order

7 transaction or while the CIT is idle waiting for a customer to

8 interact with the CIT. It will therefore be appreciated by those

9 skilled in the art that yet other modifications could be made to

10 the provided invention without deviating from its spirit and scope

11 as claimed.

32